

# **EVALUACIÓN DEL MODELO 3D DE UNA ESCULTURA OBTENIDO CON TÉCNICAS DE FOTOGRAMETRÍA A PARTIR DE IMÁGENES DE CORTO ALCANCE CAPTURADAS CON TELÉFONOS INTELIGENTES**

Estudiante:

Diana Sofía Muñoz Bedoya

Código: 1124010-3740

Trabajo presentado para optar por el título de:

Ingeniera Topográfica

PhD. Humberto Loaiza Correa

Director

Ing. Jean Pierre Diaz Paz

Director

Programa Académico de Ingeniería Topográfica

Escuela de Ingeniería Civil y Geomática

Universidad del Valle

Cali - Colombia

27 de marzo de 2017

# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Planteamiento del Problema</b>	<b>3</b>
<b>3. Objetivos</b>	<b>5</b>
3.1. Objetivo General . . . . .	5
3.1.1. Objetivos Específicos . . . . .	5
<b>4. Justificación</b>	<b>6</b>
<b>5. Marco Conceptual</b>	<b>7</b>
5.1. Visión Artificial . . . . .	7
5.2. Fotogrametría . . . . .	7
5.2.1. Fotogrametría Aérea . . . . .	8
5.2.2. Fotogrametría Terrestre . . . . .	8
5.2.3. Fotogrametría de corto alcance . . . . .	8
5.3. Cámara . . . . .	9
5.3.1. Cámara métrica . . . . .	9
5.3.2. Cámara semi-métrica . . . . .	9
5.3.3. Cámara no métrica . . . . .	9
5.4. Teléfono Inteligente . . . . .	9
5.5. Sistema Microelectromecánico . . . . .	10
<b>6. Marco Teórico</b>	<b>11</b>
6.1. Modelo de la Cámara . . . . .	11
6.1.1. Modelo Pin Hole . . . . .	11
6.2. Geometría epipolar . . . . .	12
6.2.1. Matriz Fundamental . . . . .	13
6.3. Reconstrucción Tridimensional . . . . .	13
6.4. Calibración de la Cámara . . . . .	14
6.4.1. Parámetros Intrínsecos. . . . .	14
6.4.2. Parámetros extrínsecos . . . . .	16
6.5. Structure From Motion (SfM) . . . . .	17
6.5.1. Puntos característicos . . . . .	18
6.5.2. Emparejar características . . . . .	18

6.5.3. Triangulación . . . . .	19
<b>7. Antecedentes</b>	<b>20</b>
7.1. Internacionales . . . . .	20
7.2. Nacionales . . . . .	21
<b>8. Solución Hardware</b>	<b>23</b>
8.1. Teléfono Inteligente . . . . .	24
8.2. Kinect . . . . .	24
8.3. Escultura y Monumento . . . . .	25
<b>9. Solución Software</b>	<b>26</b>
9.1. Estéreo calibración de la cámara . . . . .	26
9.2. Estimación de la matriz fundamental . . . . .	27
9.3. Rectificación de las imágenes . . . . .	28
9.4. Mapa de Disparidad . . . . .	28
9.5. Triangulación . . . . .	28
<b>10.Pruebas y Resultados</b>	<b>29</b>
10.1. Estéreo calibración de la cámara . . . . .	29
10.1.1. Procedimiento . . . . .	29
10.1.2. Resultados . . . . .	30
10.2. Mapa de Disparidad . . . . .	32
10.2.1. Procedimiento . . . . .	32
10.2.2. Resultados . . . . .	34
10.3. Reconstrucción de la Escena . . . . .	35
10.3.1. Procedimiento . . . . .	35
10.3.2. Resultados . . . . .	35
10.4. Evaluación de la nube de puntos . . . . .	36
10.5. Limitaciones . . . . .	41
<b>11.Conclusiones</b>	<b>42</b>
<b>12.Trabajos Futuros</b>	<b>44</b>
<b>Bibliografía</b>	<b>44</b>

# Índice de figuras

6.1. Modelo Pin Hole - Fuente: ( <i>Nájera</i> , 2009)	12
6.2. Modelo de geometría epipolar	13
6.3. Parámetros extrínsecos de la cámara. Fuente: <i>Rosado</i> (2012)	16
6.4. Structure from motion	17
6.5. Fases de desarrollo SfM	18
8.1. Representación del sistema	23
8.2. Plataforma física del sistema. a) Telefono inteligente b) Kinect c) Objeto a reconstruir	24
8.3. Escultura “Representaciones Humanas en Piedra”	25
9.1. Esquema metodológico	26
9.2. Plantilla de Calibración	27
10.1. Visualización de parámetros extrínsecos de la calibración. a). Vista de parámetros extrínsecos. b). Imágenes rectificadas.	30
10.2. Error de reproyección	31
10.3. Imágenes rectificadas	32
10.4. Configuración de las cámaras del par estéreo	33
10.5. Relación geométrica para obtener la profundidad Z.	33
10.6. Imágenes procesadas. a) Imagen cámara izquierda b) Imagen cámara derecha	34
10.7. Mapa de Disparidad.	34
10.8. Reconstrucción de la escena	35
10.9. Reconstrucción escultura de Moises	36
10.10 Dimensiones reales de los objetos	37
10.11 Alineación nube de puntos con modelo del Kinect	38
10.12 Representación gráfica de distancias entre nube de puntos y Kinect	39
10.13 Alineación nube de puntos con modelo de la escultura	40
10.14 Distancia entre las dos superficies	40

# Índice de cuadros

5.1. Sensores del Sistema Microelectromecánico . . . . .	10
8.1. Especificaciones técnicas de la cámara del s6 . . . . .	24
8.2. Caracterísiticas principales del Kinect . . . . .	25
10.1. Parámetros obtenidos de la estéreo calibración . . . . .	31
10.2. Muestra de distancias . . . . .	37
10.3. Puntos seleccionados para la alineación . . . . .	38
10.4. Error calculado en los puntos de referencia . . . . .	39

## Agradecimientos

*A Dios que llenó mi corazón de paciencia y perseverancia una y otra vez cuando sentía desfallecer y me brindó tranquilidad ante las situaciones que me sentí afligida.*

*A mis Padres por su inmenso esfuerzo y dedicación.*

*A mis Hermanos por ser mi inspiración.*

*A la Universidad del Valle, por brindarme los espacios que me permitieron aprender, crecer y formarme profesionalmente, donde viví gratos momentos y tuve la oportunidad de compartir con grandes seres humanos.*

*Por último y no menos importante, a mi compañero de vida, mi amigo, mi apoyo, Hanley Rivera, por su comprensión, paciencia y amor en todo momento.*

## Resumen

La reconstrucción tridimensional a partir de múltiples imágenes es un enfoque eficiente en cuanto a costo de adquisición y almacenamiento de la información. Los algoritmos para la orientación automática de imágenes son un tema importante de investigación en *Visión Artificial* y *Fotogrametría*, ya que se han desarrollado algoritmos basados en correspondencia de características y Structure From Motion (SfM), que han evolucionado a través del fortalecimiento de sus fundamentos matemáticos, estadísticos y geométricos en beneficio de mejorar su robustez y rendimiento global.

El presente trabajo de grado expone la aplicación de una metodología para la generación de un modelo tridimensional usando imágenes terrestres capturadas con la cámara de un teléfono inteligente. La metodología combina las técnicas de *Visión Artificial* y *Fotogrametría* para generar un sistema automático de reconstrucción a partir de las relaciones entre puntos característicos identificados en múltiples imágenes. Las imágenes para reconstrucción son almacenadas en el dispositivo móvil y el sistema de reconstrucción se desarrolla en el lenguaje de programación de Matlab, conformado con procesos para la identificación y registro de características, estimación de la geometría y reconstrucción de la escena.

El sistema se pone a prueba con diferentes conjuntos de imágenes con la intención de identificar las condiciones de la escena que permitan un mejor resultado en la reconstrucción, adicionalmente se realizó la integración del modelo resultante con otro de referencia obtenido con el dispositivo Kinect usando el software CloudCompare para establecer cuantitativamente la precisión de la información generada.

Las pruebas mostraron buenos resultados usando imágenes donde las condiciones de las cámaras respecto a: iluminación, sensibilidad, enfoque y balance de blancos estuvieran configuradas equivalentemente, obteniendo una precisión de 65.53 %.

**Palabras Clave:** Visión Artificial, Fotogrametría, Structure From Motion, Múltiples Vistas.

# Capítulo 1

## Introducción

Los teléfonos inteligentes son el producto de la unión de diferentes tecnologías y el continuo avance de investigación en electrónica y telecomunicaciones. [Kim \(2014\)](#) señala que los avances en los sistemas micro-electromecánicos (*Microelectromechanical Systems, MEMS*) incorporados en éstos han ido en aumento, en particular, en características como la alta resolución de la cámara, GPS, acelerómetros y magnetómetros.

[Ginovart et al. \(2014\)](#) exponen que la combinación de estas tecnologías junto con la disciplina de la *Fotogrametría* se ha convertido en una metodología de bajo costo para la adquisición de datos. En años recientes el uso de la Fotogrametría ha sido significativa en muchas áreas como topografía, arqueología, arquitectura, reconstrucción facial y de accidentes, entre otras, ya que logra identificar la geometría de cualquier objeto tridimensional, además, permite la reconstrucción de información en tres dimensiones con la cámara digital como sensor principal, ya que desde las imágenes es posible extraer las características de un objeto estimando su forma y tamaño. ([Zheng et al., 2009](#)). Entre las áreas de investigación de la Fotogrametría se encuentra la Fotogrametría Terrestre, conocida como la herramienta de medición indirecta más extendida en las labores de cartografía y de documentación patrimonial ([Lerma et al., 2011](#)), el levantamiento y reconocimiento, ya que a partir de imágenes digitales se restaura un modelo semejante a la realidad. ([Yuan-yuan et al., 2012](#)).

Este trabajo de grado hace frente a la correspondencia entre las imágenes capturadas con la cámara de los teléfonos inteligentes y la información del mundo real, teniendo en cuenta los diferentes sistemas de referencia implicados. Por lo tanto, propone un procedimiento para evaluar la calidad de los resultados obtenidos, a fin de que se logre evidenciar que la unión de nuevas tecnologías y las técnicas de *Fotogrametría* y *Visión Artificial*, permiten crear, obtener y visualizar la reconstrucción de una escena.

El contenido de éste documento está distribuido en trece capítulos. En el primero se realiza la introducción, en el segundo el planteamiento del problema de investigación, los objetivos general y específicos del trabajo se presentan en el capítulo tres; a continuación, en el capítulo cuatro se encuentra la justificación,



seguido, la búsqueda bibliográfica, organizada en un marco conceptual, marco teórico y antecedentes (capítulos cinco, seis y siete respectivamente), la metodología utilizada está estructurada en los capítulos referente al hardware y software del sistema (capítulos ocho y nueve, respectivamente); las pruebas y los resultados obtenidos están en el capítulo diez; en el capítulo undécimo se presentan las conclusiones del trabajo; y por último en los capítulos duodécimo y decimotercero, se presentan las limitaciones y los trabajos futuros.

## Capítulo 2

# Planteamiento del Problema

Según *Jauregui et al. (2013b)*, la necesidad de preservar los monumentos históricos y el patrimonio cultural requiere de una apropiada documentación, lo cual incluye desde sus características geométricas hasta los menores detalles. La técnica de fotogrametría terrestre posibilita la captura de esta información y unida por medio de una herramienta CAD con datos geométricos permite el correcto uso de ésta. (*Almagro, 2009*).

La fotogrametría y las nuevas tecnologías de sensores incorporados en los teléfonos inteligentes, permiten obtener imágenes apropiadas para los procesos de reconstrucción (*Kim et al., 2013*); para lograrlo, usualmente se requiere además de la cámara, un dispositivo GPS, una unidad de medición inercial (la cual incluye giroscopio, magnetómetro y acelerómetro), que permiten la georeferenciación de las imágenes obtenidas.

Las cámaras que tienen los teléfonos inteligentes cuentan con diferentes modos de enfoque: automático y manual, los cuales están diseñados para conseguir imágenes en rangos de corto alcance; por lo que no cuentan con las características de las imágenes usadas en fotogrametría convencional. (*Kim et al., 2013*). A pesar de las limitaciones que se pueden tener debido a que este tipo de cámara está categorizada como cámara no métrica, es decir, no está construida específicamente para propósitos fotogramétricos (*Yun et al., 2012*), la tecnología de los teléfonos inteligentes sirve como herramienta para generar modelos tridimensionales, empleando técnicas no destructivas como *Visión Artificial* que permite obtener, analizar y procesar señales digitales (*Quintero et al., 2011*).

La modelización 3D es aplicada con éxito en labores de marketing urbanístico, planificación territorial, medio ambiente, turismo y recreaciones patrimoniales (monumentos histórico – artístico, edificios singulares, esculturas, yacimientos arqueológicos) (*Lerma, 2011*).

*Rosado (2012)* implementó un sistema de visión, que a partir de la captura de una serie de imágenes con una cámara convencional es capaz de simular el efecto estereoscópico y con esto la reconstrucción tridimensional de la escena.

Una investigación llevada a cabo en la Universidad de California, evaluó la

posibilidad de utilizar un teléfono inteligente equipado con *MEMS* para realizar un seguimiento y rastreo de la realidad para luego ser comparado con sistemas fotogramétricos existentes. (*Chiabrando et al., 2011*).

El Proyecto Redentor, desarrollado por *Chen et al.* en la estatua del Cristo de Corcovado en Brasil, consistió en la adquisición de 3.584 imágenes que abarcan una a una todas las partes que lo componen, haciendo uso de puntos de control y un Software especializado que generó una nube de puntos y posteriormente la textura a partir de una malla compuesta por 2.5 millones de triángulos.

Con base en las evidencias mencionadas anteriormente y teniendo en cuenta que la conservación de estructuras y monumentos no solo se basa en la restauración, sino también en el registro, protección y promoción de las mismas, para asegurar su permanencia en condiciones originales, esta investigación se concentra en evaluar el modelo tridimensional generado desde imágenes capturadas con un teléfono inteligente, poniendo a prueba los sensores incorporados en él, para realizar la reconstrucción de una escultura. A partir de esto, surge la siguiente pregunta de investigación; ¿Cuál es la precisión del modelo digital tridimensional de una escultura obtenido a partir de imágenes capturadas con teléfonos inteligentes?

## Capítulo 3

# Objetivos

### 3.1. Objetivo General

Evaluar la precisión de un modelo geométrico 3D obtenido mediante procesos fotogramétricos, usando imágenes de una escultura capturadas con teléfonos inteligentes.

#### 3.1.1. Objetivos Específicos

- Determinar las principales metodologías para generar, evaluar y validar la precisión de un modelo tridimensional.
- Elegir una plataforma hardware y software para la captura de imágenes y datos complementarios.
- Implementar un procedimiento para generar modelos tridimensionales.
- Establecer un protocolo de pruebas para validar la calidad de los modelos tridimensionales obtenidos.

## Capítulo 4

# Justificación

Los seres humanos perciben el mundo constantemente en tres dimensiones. La ciencia y la tecnología han aprovechado esta capacidad para obtener modelos tridimensionales en interfaces gráficas de computación, las cuales corresponden a sistemas visuales artificiales, que permiten medir de forma precisa los objetos que se encuentran en el mundo real.

La extracción de estas medidas a partir de una o varias imágenes es el objetivo principal de la *Fotogrametría*, herramienta valiosa en la recopilación de información([Hernández López, 2006](#)), ya que su objetivo es obtener información métrica de los objetos que están en el entorno y mediante ésta generar modelos tridimensionales. Adicionalmente, la *Visión Artificial*, centra sus objetivos en generar algoritmos que realicen la reconstrucción de objetos en forma automática y sistemas que sean capaces de extraer información relevante de las imágenes.([Vite Silva, 2008](#)).

Teniendo en cuenta lo anterior, existe un interés creciente en la modelización tridimensional. Recientemente, se han desarrollado diferentes herramientas con el objetivo de crear dichos modelos utilizando series de fotografías obtenidas desde distintos puntos de vista, utilizando algoritmos de detección automática de puntos homólogos y deducción de parámetros de transformación que permiten hallar puntos 3D con relativa precisión([Santamaría and Sanz, 2013](#)).

La unión de las técnicas de *Fotogrametría* y *Visión Artificial*, es una contribución a los avances en reconstrucción tridimensional, pues responderá a exitosas estrategias de captura y generación de información ([Jauregui et al., 2013b](#)).

El presente trabajo de grado busca contribuir al problema de reconstrucción tridimensional, a partir de imágenes capturadas con teléfonos inteligentes, centrando principal interés en imágenes de esculturas de valor histórico y representativo, y de esta manera, aportando al progreso en la adquisición de información para la reconstrucción de geometría tridimensional de objetos a partir de imágenes digitales, dando paso a una revolución en los sistemas de generación de información tridimensional, a partir del procesamiento de la información obtenida con sensores de fácil acceso.

## Capítulo 5

# Marco Conceptual

Este marco conceptual brinda a la investigación una estructura ordenada y coherente de los conceptos que permiten abordar la reconstrucción tridimensional. Aquí se relacionan algunos métodos y modelos presentes en la literatura que hacen parte fundamental de la reconstrucción tridimensional.

### 5.1. Visión Artificial

La visión artificial describe el mundo que el ser humano vé, en una o más imágenes para reconstruir sus propiedades, tales como: forma, iluminación y distribución del color. Los investigadores en visión artificial han desarrollado en conjunto con la matemática, técnicas para la recuperación de la forma tridimensional y apariencia de los objetos a partir de imágenes. (*Szeliski, 2010*).

### 5.2. Fotogrametría

La American Society of Photogrammetry and Remote Sensing (ASPRS) define la Fotogrametría como; “... *el arte, ciencia y tecnología de obtener información fiable de los objetos físicos y del medio ambiente mediante procesos de registro, medición e interpretación de imágenes fotográficas ...*”. (*Jauregui et al., 2013b*).

Se puede describir a la fotogrametría como la técnica de medición indirecta, que a partir de imágenes (digitales o análogas) y principios matemáticos establecidos obtiene las características geométricas tridimensionales de elementos en el mundo real. (*Espín and Guano, 2013*).

Según *Hartley and Zisserman (2004)*, si se dispone de dos (o más) imágenes del mismo objeto tomadas desde puntos de vista significativamente diferentes, se puede obtener información tridimensional de cualquier punto del objeto presente en las dos imágenes.

Según *Hernández López (2006)* el desarrollo de la fotogrametría se ha llevado a cabo por técnicas diferentes que incluyen las siguientes etapas:

- Analógica

La fotogrametría analógica es básicamente la determinación de un objeto en el espacio apoyada en los equipos de restitución y con la utilización directa de fotografías aéreas para formar un modelo estereoscópico, basado en el empleo de componentes ópticos.

- Analítica

La fotogrametría analítica introduce los cálculos y restitución en tiempo real, pues enfrenta el problema mediante modelos matemáticos, incorporando la corrección de errores sistemáticos y obteniendo productos digitales.

- Digital

Los avances en informática acontecidos en la última década del siglo XX han permitido el avance de la fotogrametría hasta su fase actual, la fase digital. Han sido necesarios avances de la informática tanto en el terreno del hardware como del software. *Hernández López (2006)*.

Según (*Jauregui et al., 2013a*), a lo largo del tiempo se han desarrollado varios métodos fotogramétricos que se han ajustado a los diferentes campos de aplicación. Esto llevó a la creación de equipos específicos capaces de aplicar estas técnicas especializadas. Se distinguen tres grandes grupos dentro de la fotogrametría, los cuales se exponen a continuación.

### 5.2.1. Fotogrametría Aérea

Aquella que utiliza fotografías capturadas desde una cámara aerotransportada, lo que implica que la captura debe ser vertical. Esta técnica es aplicada para los trabajos cartográficos de amplia extensión.

### 5.2.2. Fotogrametría Terrestre

Es la técnica que utiliza fotografías tomadas sobre un soporte terrestre, lo que hace necesario conocer parámetros como posición geográfica y elementos de orientación externa de la cámara. Fue la primera aplicación práctica de la fotogrametría y hoy en día se usa principalmente en labores de apoyo a la arquitectura, arqueología y en levantamientos topográficos.

### 5.2.3. Fotogrametría de corto alcance

Agrupar aquellas aplicaciones que no tienen carácter geodésico o topográfico, y se aplica para resolver problemas muy específicos. Esta división abarca mayor amplitud de técnicas para la toma de fotografías y su posterior restitución, incluyendo aplicaciones arquitectónicas y arqueológicas.

## 5.3. Cámara

Según [Vozmediano \(2006\)](#), este dispositivo constituye uno de los elementos más importantes en cualquier sistema de captura de información visual; su función básica es convertir una escena tridimensional en una señal temporal, analógica o digital, que pueda ser tratada, almacenada o modulada para transmitirse a un receptor remoto.

Las cámaras pueden clasificarse según ciertas variables, como por ejemplo, la precisión y modo de captura. Considerando el criterio de precisión se tienen cámaras métricas, semi-métricas y no métricas.

### 5.3.1. Cámara métrica

Este tipo de cámaras permite la realización de levantamientos fotogramétricos y topográficos. La cámara métrica tiene la característica de tener una calibración muy exacta de sus parámetros ópticos, además de una mayor definición en la imagen, pues su resolución óptica es mayor que las cámaras convencionales. ([Jauregui, 2006](#))

### 5.3.2. Cámara semi-métrica

Las cámaras semi-métricas son dispositivos más sólidos, debido a la ubicación de una retícula de calibración; la cual permite eliminar hasta cierto punto la distorsión y la deformación de las imágenes. ([Lerma et al., 2011](#))

### 5.3.3. Cámara no métrica

Estas cámaras son las que priorizan la calidad radiométrica sobre la calidad geométrica, provocada por la distorsión del objetivo. En estos dispositivos no existen marcas fiduciales o retículas graduadas, por lo que se requiere una calibración previa para usos fotogramétricos. La utilización de esta es muy amplia en aplicaciones de fotogrametría terrestre y de corto alcance. ([Jauregui, 2006](#))

## 5.4. Teléfono Inteligente

El término inteligente de los teléfonos celulares, hace referencia a un terminal con teclado QWERTY en miniatura, una pantalla táctil, un administrador de aplicaciones PIM (*Personal Information Manager*), hardware para servicios multimedia y un sistema operativo que soporte el desarrollo de tareas de mediana complejidad similares a las que se realizan en un equipo informático, combinan las funciones de un teléfono celular tradicional con otras características, como la posibilidad de instalar en el dispositivo un sistema operativo completo, con aplicaciones para realizar diversas tareas similares a las de un computador portátil que permite trabajar con grandes cantidades de datos. El sistema operativo que funciona bajo ciertos teléfonos inteligentes está



definido como el conjunto de órdenes y programas que controlan los procesos básicos y permiten el funcionamiento de otros programas. (*Kim et al., 2013*)

## 5.5. Sistema Microelectromecánico

Se refiere a sensores compuestos de elementos electromecánicos a nivel microscópico, que promueven la mejora en captura y proceso de información en términos de capacidad, energía y de tiempo de respuesta (*Yun et al., 2012*). El cuadro 5.1 muestra algunos de los sensores que hacen parte del MEMS;

Sensor	Descripción
Acelerómetro	Detecta medidas de vibración ocasionando una señal eléctrica.
Giroscopio	Mide velocidad angular y brinda la orientación del dispositivo.
Magnetómetro	Mide la fuerza o la dirección de la señal magnética.

Cuadro 5.1: Sensores del Sistema Microelectromecánico

## Capítulo 6

# Marco Teórico

Este capítulo enmarca la teoría desde la que se abordará la reconstrucción tridimensional. Aquí se plantean los conceptos elementales como: modelo de la cámara, calibración de la cámara, junto con los modelos empleados para lograr el objetivo general.

### 6.1. Modelo de la Cámara

La cámara es una herramienta que proyecta objetos 3D (de la escena) en el espacio 2D (imagen de los objetos). Para representar geoméricamente tal proyección se usa el modelo Pin Hole (ver figura 6.1); en este modelo la imagen resulta de la proyección de todos los puntos de la escena sobre el plano de la imagen a través de un solo punto llamado centro de proyección. (Nájera, 2009).

#### 6.1.1. Modelo Pin Hole

El sistema de proyección de la cámara es representado geoméricamente a través del modelo de la cámara Pin Hole, el cual funciona bajo la hipótesis de que la imagen es la proyección del objeto capturado desde un mismo punto (foco). Se encuentra ligado a un sistema coordenado con origen en  $F$ . Los ejes  $(F, x)$  y  $(F, y)$  son los ejes paralelos a la dirección de las filas y las columnas de los pixeles de la imagen; el eje  $(F, z)$ , conocido como centro óptico, está orientado en dirección a la escena y es perpendicular al plano de la imagen. Éste indica que cada punto de la escena  $B = (X, Y, Z)$  pasa a través del centro óptico para proyectarse en el plano de la imagen  $b = (u, v)$ , que representa valores de filas y columnas. (Figura 6.1) Nájera (2009).

$$w[x \ y \ 1] = [X \ Y \ Z \ 1] \begin{bmatrix} R \\ t \end{bmatrix} K \quad (6.1)$$

Donde

$w$  es un factor de escala arbitrario

$R$  representa la matriz de rotación de la cámara.  
 $t$  representa la matriz de traslación de la cámara.  
 $K$  representa la matriz de parámetros intrínsecos de la cámara.

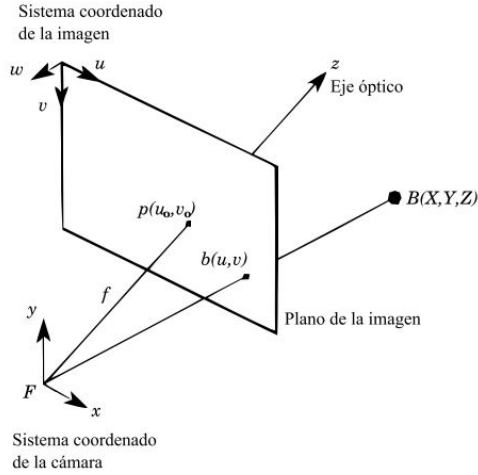


Figura 6.1: Modelo Pin Hole - Fuente: (Nájera, 2009)

Si los puntos en el mundo y los puntos en la imagen están representados por vectores homogéneos, entonces la proyección se puede expresar como una transformación lineal entre sus coordenadas homogéneas. (Sánchez et al., 2004). De esta manera, en forma matricial y utilizando coordenadas homogéneas se tiene  $B = (X, Y, Z, 1)$ ; su proyección en la imagen  $b = (u, v)$ , según el sistema de coordenadas de la cámara, está dado por la ecuación 6.2:

$$[u, v, 1] = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (6.2)$$

## 6.2. Geometría epipolar

La geometría epipolar es la geometría proyectiva intrínseca entre dos vistas. Es independiente de la estructura de la escena, y sólo depende de los parámetros internos de las cámaras y su posición relativa. (Ma et al., 2006).

La Figura 6.2 indica cómo está compuesto un plano epipolar;  $\mathbf{X}$  es un punto tridimensional de la escena,  $x$  y  $x'$  representan su proyección en el plano de la imagen y  $C$  y  $C'$  indican los centros de las cámaras. (Bejarano, 2013).

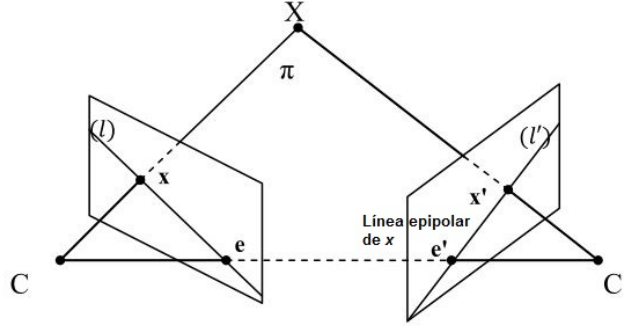


Figura 6.2: Modelo de geometría epipolar

### 6.2.1. Matriz Fundamental

La matriz fundamental  $\mathbf{F}$  es la representación algebraica de la geometría epipolar, se deriva desde la correspondencia entre un punto y su línea epipolar. La figura 6.2 expone que dado un par de imágenes, para cada punto  $x$  en una imagen existe una línea epipolar  $l'$  correspondiente en la otra imagen, cualquier punto  $x'$  en la segunda imagen que coincida con el punto  $x$  debe estar en la línea epipolar  $l'$ . (Rosado, 2012). La línea epipolar es la proyección en la segunda imagen del rayo desde el punto  $x$  a través del centro de la cámara. Por lo tanto, hay una relación desde un punto en una imagen para la correspondencia de una línea epipolar en la otra imagen.  $x \rightarrow l'$  (Bejarano, 2013).

Una de las propiedades más importantes de la matriz fundamental es la condición de correspondencia, que establece que para cualquier correspondencia de puntos  $x \leftrightarrow x'$  en dos imágenes se satisface la condición de la ecuación 6.3:

$$x'^T \mathbf{F} x = 0 \quad (6.3)$$

Esta condición se cumple dado que, si  $x$  y  $x'$  son correspondientes,  $x'$  yace sobre la línea epipolar  $l' = \mathbf{F}x$  que corresponde al punto  $x$ , de tal forma que  $0 = x'^T l' = x'^T \mathbf{F}x$ . De forma inversa, si los puntos en la imagen satisfacen la relación 6.3, entonces las líneas definidas por estos puntos son coplanares. Esta es una condición necesaria para determinar que los puntos son correspondientes. (Vite Silva, 2008).

## 6.3. Reconstrucción Tridimensional

Los métodos que componen la fotogrametría están fundamentados en la estereoscopia, la cual formula que a partir de dos o más imágenes de un mismo objeto es posible calcular la posición 3D de éste. Por esta razón, es necesario conocer la correspondencia de puntos  $x \leftrightarrow x'$  entre las imágenes, donde  $x$

pertenece al conjunto de puntos de una primera imagen y  $x'$  pertenece al conjunto de puntos de una segunda imagen. ([Rosado, 2012](#)).

La reconstrucción tridimensional está compuesta por los siguientes procesos:

- **Calibración de la cámara**

En esta fase se hallan los modelos matemáticos y los parámetros que representan la proyección del objeto en la formación de las imágenes, una vez calculados los parámetros intrínsecos y extrínsecos de la cámara se continua con la fase de pre-procesamiento.

- **Pre-procesamiento**

La fase de pre-procesamiento está compuesta por la rectificación de las imágenes y permite obtener las líneas epipolares entre éstas. Una vez generadas las líneas epipolares se inicia la fase de correspondencia.

- **Correspondencia**

Esta fase es donde se busca identificar los pixeles que corresponden al objeto del mundo real en las imágenes.

- **Triangulación**

Finalmente la fase de **Triangulación**, es donde se hallan las coordenadas de cada punto emparejado en la fase anterior.

## 6.4. Calibración de la Cámara

El proceso de calibración de la cámara permite conocer el valor de los parámetros intrínsecos y extrínsecos tales como: distancia focal, factores de distorsión, puntos centrales del plano de la imagen, rotación y traslación, respectivamente. ([Sun et al., 2015](#)).

Para conocer los parámetros nombrados anteriormente, deben ser calculados los coeficientes de la matriz de proyección, la cual establece los puntos en la mira de calibración y la respectiva proyección en la imagen.

### 6.4.1. Parámetros Intrínsecos.

Son aquellos que definen el modelo de la cámara, estableciendo sus características geométricas y ópticas. ([Gonzalez, 2003](#)).

Los parámetros intrínsecos son obtenidos a partir del producto de la matriz  $K$  y la matriz  $P$ , las cuales son la transformación del dispositivo al plano de la imagen y la proyección perspectiva, respectivamente.

$$K = \begin{bmatrix} -k_u & 0 & 0 & u_0 \\ 0 & k_v & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6.4)$$

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{f} & 0 \end{bmatrix} \quad (6.5)$$

La ecuación 6.4 está conformada por los parámetros  $u_0, v_0$  que definen las coordenadas del foco y  $k_u$  y  $k_v$  que determinan los factores de escala horizontal y vertical respectivamente.

$$KP = \begin{bmatrix} k_u & 0 & 0 & u_0 \\ 0 & -k_v & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{f} & 0 \end{bmatrix} = \begin{bmatrix} k_u & 0 & \frac{u_0}{f} & 0 \\ 0 & -k_v & \frac{v_0}{f} & 0 \\ 0 & 0 & \frac{1}{f} & 0 \end{bmatrix} \quad (6.6)$$

El resultado obtenido de  $KP$  es multiplicado por el factor común  $f$ , obteniendo la ecuación 6.7, que representa una aplicación lineal del espacio proyectivo. (*Hartley and Zisserman, 2004*).

$$I_c = \begin{bmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (6.7)$$

Donde:

$\alpha_u$ , es el factor de escala horizontal multiplicado por el factor  $f$ :  $\alpha_u = k_u f$ .

$\alpha_v$ , es el factor de escala vertical multiplicado por el factor  $f$ :  $\alpha_v = k_v f$ .

Obteniendo las coordenadas de la cámara sin sus dimensiones en  $x_c, y_c, z_c$ , se obtiene que  $x_c = x/z$ ,  $y_c = y/z$  y  $z_c = 1$ , lo que permite describir la relación entre el sistema de coordenadas de la imagen y la cámara con las ecuaciones 6.8 y 6.9, que pueden ser representadas en estructura matricial como lo muestra la ecuación 6.10.

$$u = \alpha_u x_c + u_0 \text{ con } \alpha_u < 0 \quad (6.8)$$

$$v = \alpha_v x_c + v_0 \quad (6.9)$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = C \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}$$

donde,

$$C = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.10)$$

De tal forma, la matriz  $I_c$  está conformada por una transformación afín entre la cámara y las imágenes,  $C$ , 6.10, y una transformación proyectiva que se presenta en la ecuación 6.11.

$$I_c = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (6.11)$$

#### 6.4.2. Parámetros extrínsecos

Los parámetros extrínsecos son los que definen la relación (rotación y traslación) entre un sistema de coordenadas absoluto y otro ligado a la cámara. (Gonzalez, 2003). En la figura 6.3 se representa gráficamente la especificación de los parámetros extrínsecos de la cámara.

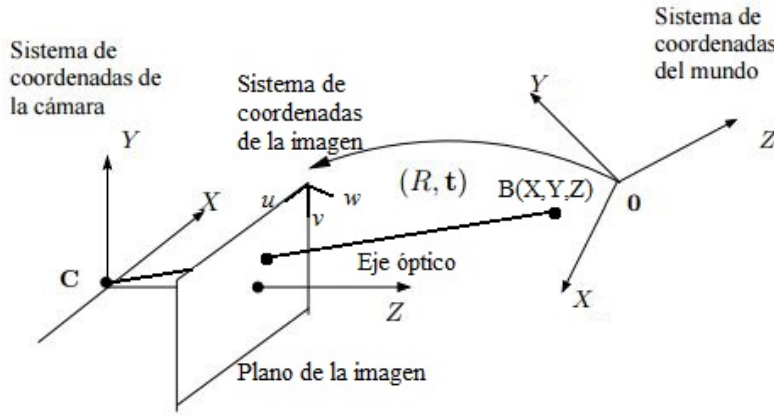


Figura 6.3: Parámetros extrínsecos de la cámara. Fuente: Rosado (2012)

La matriz de parámetros extrínsecos está compuesta por un vector de traslación  $t$ , que traslada el origen del sistema de coordenadas del objeto en el espacio al sistema de coordenadas de la cámara y una matriz de rotación  $R$ , que alinea en la misma dirección y sentido los ejes  $(X,Y,Z)$  de los sistemas de coordenadas. (Rosado, 2012). Estos parámetros se pueden hallar por medio de dos procesos de transformación y uno de proyección, que hacen referencia a la mira de calibración - imagen y dispositivo de captura - imagen, donde la primera se realiza a través de la ecuación 6.12. (Chiabrando et al., 2011)

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (6.12)$$

Donde:

$x, y, z$ , sistema coordenado de la imagen.

$X, Y, Z$ , sistema coordinado de la mira de calibración.  
 $r$ , matriz de rotación.  
 $t$ , vector de traslaciones.

## 6.5. Structure From Motion (SfM)

La técnica de *Structure from Motion* se basa en la reconstrucción de la escena a partir del movimiento de la misma. Opera bajo los mismos principios de *Fotogrametría Estereoscópica*, es decir, que una estructura 3D puede ser resuelta a partir de una serie de imágenes superpuestas. (Figura 6.4). Sin embargo, difiere fundamentalmente de la fotogrametría convencional, en que la geometría de la escena, posición y orientación de la cámara se resuelve automáticamente, sin la necesidad de especificar a priori una red de objetivos de los que se conocen las posiciones 3D. (*Westoby et al.*, 2012).

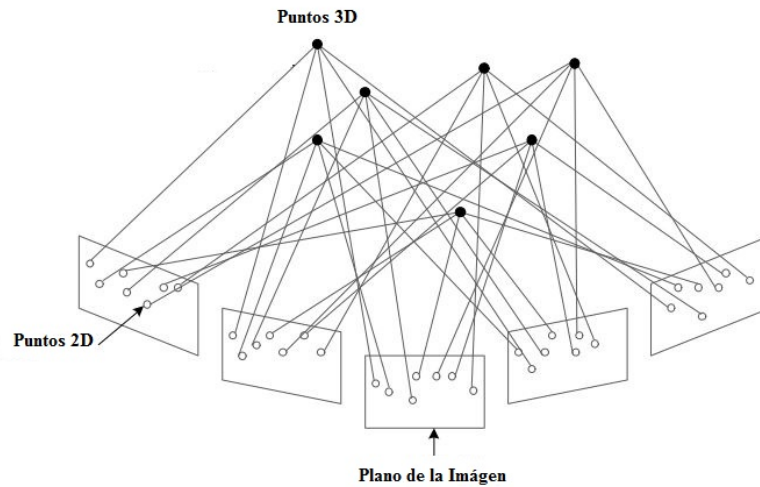


Figura 6.4: Structure from motion

Autores como *Roncella et al.* (2011), *Szeliski* (2010), *Westoby et al.* (2012) plantean las siguientes fases para llevarlo a cabo (Figura 6.5).



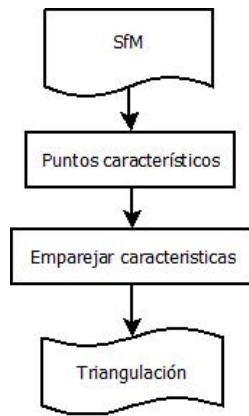


Figura 6.5: Fases de desarrollo SfM

#### 6.5.1. Puntos característicos

El primer paso consiste en la identificación de puntos de interés o puntos característicos. Uno de los más importantes requerimientos para identificar los puntos característicos es que estos se puedan apreciar en la imagen consecutiva; si esto no se cumple, no sería posible la correspondencia singularmente con los puntos correspondientes en otra imagen. (*Bay et al., 2008*).

El algoritmo Speeded Up Robust Features (SURF como lo indican sus siglas en inglés) está basado en propiedades similares al descriptor Scale invariant feature transform (SIFT) y presenta principalmente mejoras en cuanto a la velocidad de cálculo sin ocasionar pérdida del rendimiento y mayor robustez ante posibles transformaciones de la imagen. Permitiendo obtener puntos especiales de una imagen basado en Haar wavelet, ya que obtiene un descriptor por cada punto encontrado. (*Bay et al., 2006*).

#### 6.5.2. Emparejar características

Las regiones de las imágenes son típicamente comparadas usando *Suma de Diferencias al Cuadrado (SSD)* o *Suma de las Diferencias Absolutas (SAD)*. Esta diferencia permite llevar a cabo la correspondencia entre características mediante la distancia euclidiana a los puntos de interés.

##### ■ Suma de Diferencias al Cuadrado (SSD)

Esta técnica de correspondencia está basada en la correlación de los píxeles para medir similitudes. Se trabaja con una ventana  $f$  de tamaño  $n \times m$  alrededor del píxel central en una de las imágenes. Ésta ventana se busca a lo largo de la línea epipolar de la imagen tomada por la otra cámara usando una ventana  $g$  de tamaño  $n \times m$ . De esta manera, SSD es definida como la distancia al cuadrado  $L^2$  de la diferencia de las ventanas  $f$  y  $g$ .

$$SSD(f, g) = ||f - g||^2 \quad (6.13)$$

El uso de la norma  $L^2$  genera que el procesamiento sea más lento en comparación a la técnica SAD, debido a que se hace una multiplicación adicional por cada pareja de píxeles comparada.

SSD con normalización exponencial es la medida óptima para la consistencia de imágenes si  $f$  y  $g$  solo se diferencian por ruido Gaussiano con una desviación estándar  $\sigma$ . (*Furukawa and Hernández, 2013*)

$$SSD(f, g) = e^{-\frac{||f-g||^2}{\sigma^2}} \epsilon[0, 1] \quad (6.14)$$

#### ■ Suma de Diferencias Absolutas (SAD)

La suma de diferencias absolutas (SAD) es muy similar a SSD, consiste en seleccionar una ventana  $f$  de tamaño  $n \times m$  alrededor del píxel que se pretende identificar en la primera imagen. Aquí, se calcula la suma de las intensidades de todos los píxeles de la ventana.

En la segunda imagen, para cada píxel de la ventana  $f$  que se encuentre en la línea epipolar, se genera una ventana  $g$  de tamaño  $n \times m$ ; de esta manera, se suman los valores de las intensidades de todos los píxeles de la ventana. En seguida se calcula el valor absoluto de la diferencia entre la suma de intensidades de la ventana  $f$  y la de cada uno de los píxeles hallados en línea epipolar de la segunda imagen a través de la ventana  $g$ .

$$SAD(f, g) = ||f - g||_1 \quad (6.15)$$

De manera similar a SSD es sensible al ruido, por lo que rara vez se utiliza en algoritmos que responden a las imágenes con una amplia variación en la iluminación. Sin embargo, es una buena medida para aplicaciones donde se garantiza las condiciones de captura para las imágenes, por ejemplo: aplicaciones en tiempo real o aplicaciones móviles. (*Furukawa and Hernández, 2013*).

### 6.5.3. Triangulación

En el proceso de reconstrucción tridimensional la triangulación tiene por objetivo recuperar los puntos tridimensionales a partir de las matrices de proyección y la correspondencia de puntos. (*Vite Silva, 2008*).

Consiste en calcular los rayos de retro-proyección de dos puntos que forman una correspondencia, el punto donde coincidan estos dos rayos formará parte del modelo tridimensional. Para cada par de correspondencias ( $x_i \leftrightarrow x'_i$ ) se puede calcular un punto 3D. (*Hartley and Zisserman, 2004*)

La figura 6.2 muestra la triangulación de un punto tridimensional a partir de la correspondencia de puntos y las matrices de proyección.

## Capítulo 7

# Antecedentes

### 7.1. Internacionales

Las investigaciones realizadas sobre generación de información tridimensional a partir de imágenes han permitido su avance por medio de diferentes métodos. En seguida se presentan algunas de las investigaciones encontradas en la literatura.

*Hu et al. (2008)* implementaron un algoritmo de reconstrucción 3D usando una secuencia de imágenes capturadas con una cámara de mano, utilizando el algoritmo SIFT para identificar las características entre imágenes; las características fueron emparejadas a través de la distancia euclidiana. La matriz fundamental les permitió obtener una reconstrucción proyectiva.

Más tarde, *Esteban et al. (2010)* desarrollaron la herramienta FIT3D, que tiene como objetivo la integración y distribución de un conjunto de herramientas que permiten obtener un modelo tridimensional completo a partir de un conjunto de imágenes calibradas. Esta fue desarrollada en el lenguaje programación de **Matlab**.

*Agarwal et al. (2011)* presentaron un sistema que logra emparejar y reconstruir escenas tridimensionales a partir de grandes colecciones de fotografías de la ciudad de Roma. El sistema que utilizaron está constituido por un conjunto de algoritmos de Emparejamiento y Reconstrucción. Los resultados experimentales obtenidos demuestran que es posible reconstruir la ciudad con 150.000 imágenes.

En el 2014 *Mccann* implementó la técnica de *Structure From Motion* (SfM) a través del lenguaje de programación **OpenCV**. El esquema metodológico general planteado fue el siguiente:

- Estimar la forma y el movimiento para una transformación proyectiva usando un método algebraico o un método de factorización.
- Convertir la transformación proyectiva a métrica.
- Calcular la matriz esencial  $E$  usando **RANSAC**.

- Calcular las matrices de la cámara.
- Calcular las posiciones 3D usando *Triangulación*

Después, *To et al. (2015)* en el campo del Patrimonio Cultural implementaron un sistema de fotogrametría, basado en las técnicas de *Structure From Motion*, ya que consideraron que es un sistema de bajo costo, gran campo de visión fácil, rápido y completo para la generación de información 3D métrica a partir de imágenes. La documentación arqueológica se hizo con una cámara digital Cannon EOS 550D del templo “One Pilla” ubicado en Hanoi, Vietnam. Los resultados obtenidos representan un modelo 3D escalado.

Seguidamente, *Chen et al. (2015)* realizaron el “*Projeto Redentor*” ejecutado por el equipo del software **Pix4D** en conjunto con **Aeryon Labs Inc.** y la **Universidad de Rio de Janeiro** para crear y mostrar la factibilidad, precisión y eficiencia del modelo 3D basado en imágenes. Usaron un Vehículo Aéreo no Tripulado (UAVs como lo indican sus siglas en inglés) para la adquisición de los datos.

## 7.2. Nacionales

En Colombia, en la **Universidad Industrial de Santander** desarrollaron una metodología para la reconstrucción tridimensional de superficies de platinas metálicas deformadas. Utilizaron técnicas no destructivas a partir del tratamiento de imágenes, que permitió obtener medidas de profundidad, área y longitud. Los resultados fueron comparados con medidas obtenidas con elementos manuales. (*Quintero et al., 2011*).

**El Grupo de Óptica y Láser** de la **Universidad del Cauca**, desarrolló un sistema en visión artificial estéreo que conlleva a obtener coordenadas 3D de una escena real con procesamiento en línea. Usaron algoritmos estéreo diseñados en **OpenCV** y soportados en el sistema operativo **Linux**. Los resultados obtenidos son presentados por nubes de puntos en un entorno gráfico virtual. (*Revelo-Luna et al., 2012*)

**El Centro de Investigaciones** de la **Universidad de Los Andes** presentó una metodología orientada a la conservación del patrimonio cultural para la visualización, modelado y reconstrucción tridimensional de la estatua *María Lionza*, cuyas características geométricas se obtuvieron mediante técnicas de fotogrametría terrestre. (*Jauregui et al., 2013b*).

**La Universidad Católica de Colombia** en unión con la **Universidad de Oviedo** y la **Universidad Internacional de La Rioja**, desarrolló una propuesta compuesta por diez fases combinando métodos pasivos y activos, usando imágenes y un láser escáner para complementar la información perdida y obtener mejores resultados en la reconstrucción de un objeto 3D. La metodología propuesta brindó eficientes resultados en objetos complejos. (*Pacheco et al., 2014*)

En la **Universidad Militar Nueva Granada** se desarrolló un trabajo que presentó el progreso que han experimentado las técnicas de reconstrucción

tridimensional de terrenos, desde las que involucran el uso de imágenes hasta las técnicas modernas que cuentan con equipos de última generación (láser escáner 2D y 3D).(*Penna S et al., 2014*).

En la **Universidad Distrital Francisco José de Caldas** emplearon una metodología para reconstrucción 3D conformada por un software especializado en fotogrametría digital de corto rango a partir de un conjunto de imágenes que posteriormente son representadas en una nube de puntos, para luego ser convertida en una malla que más tarde fue revestida con una textura obtenida a partir de las imágenes utilizadas. (*Fernandez et al., 2015*)

## Capítulo 8

# Solución Hardware

Este capítulo corresponde a la descripción de los componentes que intervienen en la reconstrucción tridimensional. Parte importante de éstos es la plataforma desarrollada para la captura de las imágenes (figura 8.1) que está conformada por un trípode que sostiene una ligera tabla donde dos apoyos sujetan los equipos de captura (teléfonos inteligentes), un equipo kinect y las figuras geométricas básicas.

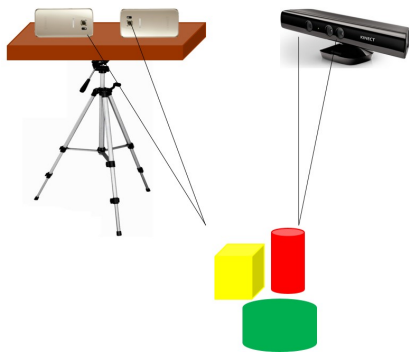


Figura 8.1: Representación del sistema

La figura 8.2 presenta la configuración real que se utilizó.



Figura 8.2: Plataforma física del sistema. a) Telefono inteligente b) Kinect c) Objeto a reconstruir

## 8.1. Teléfono Inteligente

El teléfono inteligente que se empleó como dispositivo de captura de las imágenes fue el Samsung Galaxy, seleccionado por las características de su cámara y por la tecnología que lo compone. El cuadro 8.1 presenta las características de las cámaras para la referencia s6 y s4 mini.

Referencia	Ítem	Descripción
s6	<b>Pantalla</b>	SuperAMOLED 5.1"
	<b>Resolución</b>	1440 x 2560 pixeles (577 ppp)
	<b>Flash</b>	32/64/128 GB (no expandible)
	<b>Cam. Principal</b>	16 MP (f.19)
	<b>Cam. Frontal</b>	
s4 mini	<b>Pantalla</b>	SuperAMOLED 4.3"
	<b>Resolución</b>	3264 x 2448 pixeles
	<b>Flash</b>	LED
	<b>Cam. Principal</b>	8 MP (f1.9)
	<b>Cam. Frontal</b>	1.9 MP

Cuadro 8.1: Especificaciones técnicas de la cámara del s6

## 8.2. Kinect

El dispositivo Kinect es desarrollado por Microsoft es utilizado en éste trabajo de grado en el prototipo desarrollado para la captura de la imagen de profundidad del modelo 3D. Con éste se evaluará el modelo obtenido a partir del teléfono inteligente.

El Kinect utilizado es versión 2.0 y las características se presentan en el cuadro 8.2.

Características Principales	
Campo de visión	Horizontal: 57°
	Vertical: 43°
	Rango de inclinación física: +/- 27°
	Rango de profundidad del sensor: 1.2 - 3.5 m
Data streams	Sensor de profundidad: 320x240 a 16 bits y 30 fps
	Cámara de color: 640x480 32-bits de color y 30 fps
	Micrófono: Audio de 16-bit a 16 kHz
Sistema de seguimiento	Hasta 6 personas
	20 articulaciones por persona
	Capacidad para mapear

Cuadro 8.2: Características principales del Kinect

### 8.3. Escultura y Monumento

El trabajo de grado se realiza con imágenes de esculturas de valor histórico, entre éstas se destaca la escultura seleccionada que se encuentra en la Ciudad Universitaria Melendez (Figura 8.3), que hace parte de la exposición “*El valle del Cauca Prehispánico: una historia por conocer*” que ofrece el Museo Arqueológico Julio C. Cubillos (MAJCC) ubicado en la Universidad del Valle.

Esta escultura es seleccionada para facilitar la identificación de sus partes en el modelo de profundidad, ya que las condiciones de iluminación en su ubicación, y también las texturas de color presentes, lo permiten.



Figura 8.3: Escultura “Representaciones Humanas en Piedra”



## Capítulo 9

# Solución Software

El proceso de la reconstrucción tridimensional a partir de imágenes involucra tres fases fundamentales: la calibración de la cámara, la estimación de los parámetros de orientación y posición y la estimación del mapa de disparidad. A partir de las cuales se han definido cinco pasos para obtener la información tridimensional: estéreo calibración de la cámara, rectificación de las imágenes, generación de mapa de disparidad, correlación entre características y finalmente reconstrucción de la escena. La figura 9.1 presenta el esquema metodológico general.

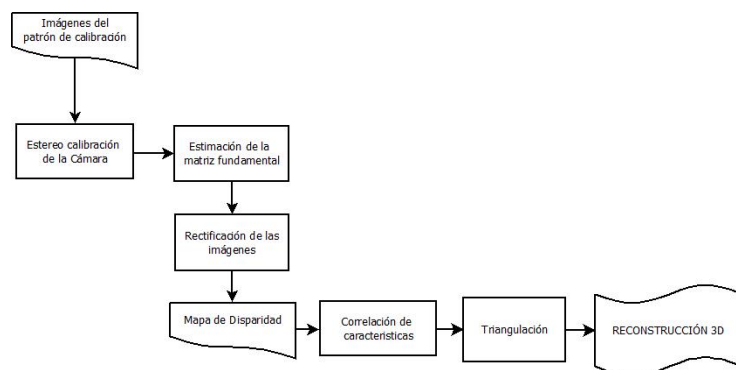


Figura 9.1: Esquema metodológico

### 9.1. Estéreo calibración de la cámara

Es necesario calibrar las cámaras y estimar los parámetros de orientación y posición de las mismas por cada imagen, con la finalidad de generar para cada una, las respectivas matrices de proyección. La calibración de la cámara se hace mediante el uso de un plantilla plana de calibración como se muestra en la figura 9.2.

Se capturaron veinticinco imágenes de la plantilla en diferentes ángulos de forma que la cuadrícula sea visualizada completamente. Las coordenadas de dichos puntos y de sus proyecciones permiten plantear un sistema de ecuaciones, cuya solución proporciona los parámetros intermedios necesarios para la estimación de los parámetros intrínsecos y extrínsecos, logrando describir los parámetros que modelan el paso de los rayos a través del lente y su proyección sobre el plano de la imagen.

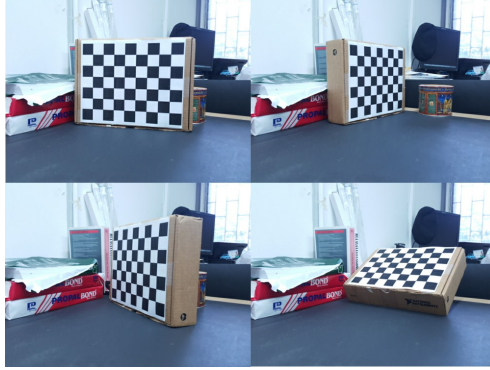


Figura 9.2: Plantilla de Calibración

## 9.2. Estimación de la matriz fundamental

Este proceso se llevó a cabo con la aplicación de calibración estéreo de la cámara de Matlab que usa los métodos descritos por [Hartley and Zisserman \(2004\)](#), donde se plantea que la matriz fundamental se puede derivar algebraicamente de la forma:

$$X(\lambda) = M^+x + \lambda\mathbf{C} \quad (9.1)$$

donde  $M^+$  es la pseudo-inversa de la matriz de proyección  $M$ , es decir,  $MM^+ = I$  (matriz identidad de orden 3 x 3),  $\mathbf{C}$  es el centro óptico de la primera imagen, y  $\lambda$  es el valor escalar para parametrizar el rayo. Dos puntos en particular del rayo son:  $M^+x$  si  $\lambda = 0$ , y el primer centro óptico  $\mathbf{C}$  si  $\lambda = \infty$ . Estos dos puntos son proyectados en la segunda imagen por la segunda matriz de proyección  $M'$  como  $M'M^+x$  y  $M'\mathbf{C}$ , respectivamente. La línea epipolar es la línea que une a este par de puntos proyectados, y se expresa como  $l' = (M'\mathbf{C})x(M'M^+x)$ . El punto  $M'\mathbf{C}$  es el epipolo  $e'$  en la segunda imagen, entonces la línea epipolar queda expresada como  $l' = [e']_x(M'M^+)x = Fx$ , donde  $[e']_x$  es la matriz antisimétrica del epipolo  $e' = [e'_1, e'_2, e'_3]^T$  expresada como,

$$[e']_x = \begin{bmatrix} 0 & -e'_3 & e'_2 \\ e'_3 & 0 & -e'_1 \\ -e'_2 & e'_1 & 0 \end{bmatrix} \quad (9.2)$$

de tal forma que la expresión que representa a la matriz fundamental es

$$F = [e']_x (M' M^+) \quad (9.3)$$

Al emplear la forma canónica en las matrices de proyección, es decir, la primera matriz de proyección  $M$  está en el origen del sistema de coordenadas del mundo  $M = K[I | \mathbf{0}]$  y la segunda matriz de proyección se expresa como  $M' = K'[R | t]$  con respecto a la primera matriz de proyección; la pseudo-inversa de  $M$  y el centro de proyección quedan expresados como,

$$M^+ = \begin{bmatrix} K^{-1} \\ 0^T \end{bmatrix} \quad \mathbf{C} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

y los epipolos se definen como,

$$\mathbf{e} = M \begin{pmatrix} -R^T t \\ 1 \end{pmatrix} = K R^T t \quad \mathbf{e}' = M' \begin{pmatrix} 0 \\ 1 \end{pmatrix} = K' t$$

### 9.3. Rectificación de las imágenes

La rectificación consiste en transformar las imágenes del par estéreo haciendo que las líneas epipolares coincidan con las filas de la imagen. De esta forma, la proyección de un punto sobre los planos de la imagen de cada cámara del par estéreo se ubica sobre la misma fila en cada imagen. El algoritmo de rectificación a partir de la matriz fundamental y la homografía del sistema calcula las nuevas coordenadas que debe ocupar cada pixel en las imágenes. ([Lin et al., 2010](#))

### 9.4. Mapa de Disparidad

La información de profundidad de un punto está determinada por la diferencia de posición en los dos planos de la imagen, cada punto emparejado en las imágenes conlleva al mapa de disparidad. Esto genera una imagen en la cual el nivel de gris de cada pixel representa la profundidad medida desde la cámara. [Rosado \(2012\)](#)

### 9.5. Triangulación

La triangulación determina la posición del punto  $X$  en el espacio físico usando el triangulo formado entre el punto observado  $P$  y los ejes ópticos de las cámaras, separados a una distancia  $B$  también conocida como línea base. ([Olsson et al., 2006](#)).

## Capítulo 10

# Pruebas y Resultados

En este capítulo se presentan los parámetros calculados en la estéreo calibración de la cámara, que son parte esencial del tratamiento de las imágenes, y seguidamente se evidencia la información extraída de las imágenes mediante el conjunto de procesos que conforman la técnica de *Structure From Motion (SfM)*. Por último se presentan los modelos obtenidos en las pruebas del sistema y la evaluación del modelo final con el modelo obtenido con el Kinect como referencia.

### 10.1. Estéreo calibración de la cámara

La estéreo calibración es el cálculo de los parámetros intrínsecos de cada cámara, además de la rotación y traslación de la segunda cámara respecto de la primera, parámetros que son posteriormente empleados para rectificar las imágenes. De esta manera, las imágenes rectificadas se utilizan para calcular el mapa de disparidad indispensable para la reconstrucción de la escena 3D.

#### 10.1.1. Procedimiento

Se diseña un sistema para visión estéreo compuesto por dos cámaras separadas a una distancia establecida de 10 cm, con el objetivo de obtener dos puntos de vista diferentes de una misma escena, semejante a la visión binocular humana.

Se realizó la captura de veinticinco imágenes del patrón de calibración, el cual consta de 63 cuadrados, distribuidos en 7 filas y 6 columnas, cada uno de 3x3 cm.

Las imágenes del patrón de calibración fueron capturadas desde ángulos y posiciones arbitrarias a una distancia aproximada de 50 cm de las cámaras. La figura 10.1 hace referencia a la interfaz de la herramienta de Calibración del Software Matlab utilizada, que muestra, en el costado izquierdo, los pares de imágenes capturados por las cámaras; al costado derecho se visualiza la rectificación de las imágenes resaltando las líneas epipolares entre éstas y en

el centro las posiciones del patrón de calibración respecto al punto fijo de las cámaras.

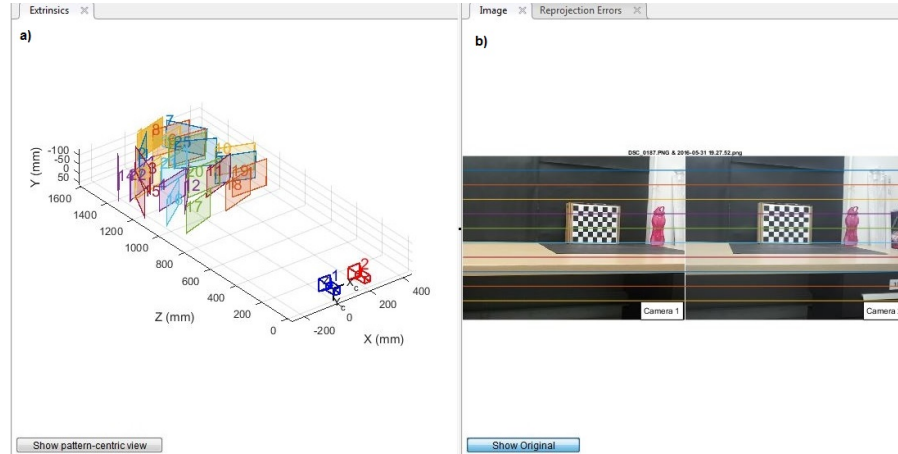


Figura 10.1: Visualización de parámetros extrínsecos de la calibración. a). Vista de parámetros extrínsecos. b). Imágenes rectificadas.

Los parámetros obtenidos del modelo son los siguientes: longitud focal, coeficientes de distorsión, error por pixel, valores del punto principal, vectores de traslación y matrices de rotación. A través del error medio del pixel se define la confiabilidad de los parámetros obtenidos, con los que posteriormente se calculan los parámetros extrínsecos de la cámara, que facilitan estimar la posición respecto al mundo real.

### 10.1.2. Resultados

La matriz K indica el desplazamiento del centro de coordenadas del plano de la imagen respecto al punto principal, el grado de perpendicularidad de las paredes de los píxeles del sensor y las distancias focales en píxeles. El cuadro 10.1 indica los parámetros obtenidos restantes.

$$K_{cam1} = \begin{bmatrix} 763.8728 & 0 & 0 \\ 0 & 767.1316 & 0 \\ 314.5963 & 229.0661 & 1 \end{bmatrix}$$

$$K_{cam2} = \begin{bmatrix} 634.7263 & 0 & 0 \\ 0 & 637.5824 & 0 \\ 314.8480 & 238.0668 & 1 \end{bmatrix}$$

La reproyección del error son las distancias calculadas en píxeles entre los puntos identificados y los puntos reproyectados del patrón de calibración. La aplicación de Matlab de Calibración Estéreo de la Cámara, calcula la reproyección del error mediante la proyección de los puntos del patrón de

Parámetro	Resultado Cam1	Resultado Cam2
Longitud focal	[763,87-767,16] mm	[763,87-767,16] mm
Punto principal	[314,59-229,06] pixeles	[314,84-238,06] pixeles
Distorsión radial	[0,03-0,18] mm	[0,17-0,23] mm

Cuadro 10.1: Parámetros obtenidos de la estéreo calibración

calibración a partir de las coordenadas del mundo, interpretados en coordenadas de la imagen.

Se visualiza en una gráfica de barras (10.2), que muestra el error de reproyección promedio por imagen a lo largo del error medio global.

El error medio por pixel estimado es:

$$prom\ error = 0.23\ pixeles$$

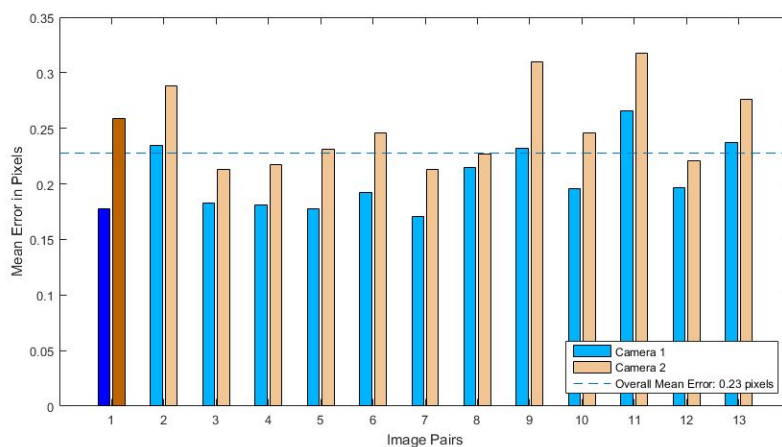


Figura 10.2: Error de reproyección

La Figura 10.3 hace referencia a la interfaz de la herramienta de Calibración Estéreo de la Cámara, dónde se muestra la recfiticación de las imágenes, éstas garantizan la precisión de la calibración puesto que las imágenes no están distorsionadas y las filas se encuentran alineadas.

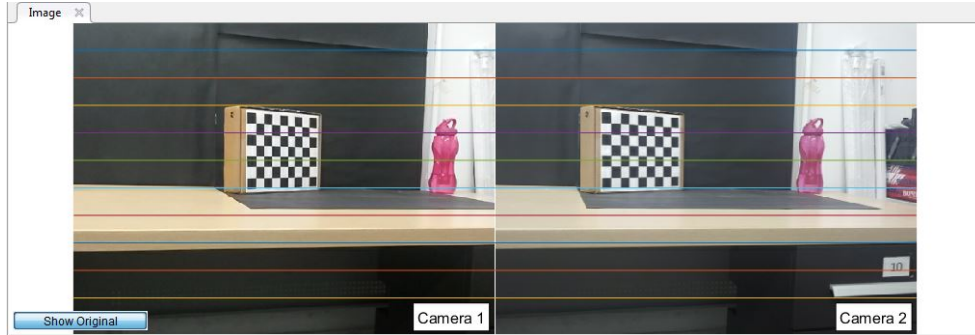


Figura 10.3: Imágenes rectificadas

## 10.2. Mapa de Disparidad

La tarea de extraer la profundidad de un conjunto de imágenes se convierte en el proceso de estimar el mapa de disparidad  $d$ , a partir del proceso de rectificación. El mapa de disparidad puede ser obtenido por la similitud de los píxeles en las ubicaciones correspondientes  $(x, y)$  y  $(x', y') = (x + d, y)$ . (Szeliski, 2010)

Siempre y cuando la posición de los centros de proyección, longitud focal y orientación de los ejes ópticos sean conocidos, la profundidad puede ser establecida por la triangulación de las disparidades obtenidas a partir del proceso de correspondencia.

Para obtener el mapa de disparidad es necesario contar con los siguientes parámetros:

1. Dos imágenes estéreo rectificadas
2. Tamaño de la ventana usado para el emparejamiento
3. Valores mínimos y máximos de disparidad

### 10.2.1. Procedimiento

Para definir la disparidad se establece una configuración de dos cámaras como se indica en la figura 10.4.

Donde:  $\overrightarrow{O_R O_R} \parallel \overrightarrow{O_L O_L}$ , son los ejes ópticos paralelos de las cámaras,  $f$  representa la distancia focal con centros en  $O_L$  y  $O_R$  separados por una distancia  $B$  (línea base), de forma que las imágenes  $I_L$  e  $I_R$  formadas están en planos paralelos. De esa manera la línea base es paralela a la coordenada  $x$  de las imágenes.

Las cámaras utilizadas cumplen con el modelo *pin hole* considerado anteriormente, donde un punto en el espacio tridimensional  $P = (X, Y, Z)^T$ , se proyecta en cada una de las imágenes bidimensionales en los puntos  $p_L$  y  $p_R$ , con coordenadas  $(x_L, y_L)^T$  y  $(x_R, y_R)^T$ , respectivamente.

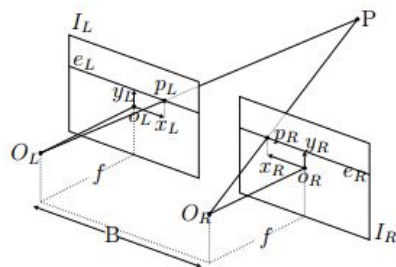


Figura 10.4: Configuración de las cámaras del par estéreo

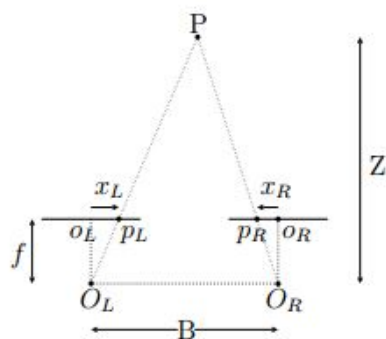


Figura 10.5: Relación geométrica para obtener la profundidad  $Z$ .

El plano que abarca a los puntos  $P$ ,  $O_L$  y  $O_R$ , une a las imágenes en dos rectas  $e_L$  y  $e_R$ . Dada la configuración del par estéreo, éstas son *rectas epipolares* entre sí, quiere decir que, un punto  $p_L$  en la recta  $e_L$  de la imagen  $I_L$  tiene su correspondiente en algún punto de la recta  $e_R$ .

La figura 10.5 muestra la correspondencia de los parámetros definidos en el par estéreo que permiten obtener la relación entre la disparidad  $d$  y la profundidad  $Z$  del punto  $P$ .

Así, la disparidad es la diferencia en las coordenadas horizontales de los puntos  $p_L$  y  $p_R$ , es decir,  $d = x_L - x_R$ . Las coordenadas de  $p_L$  y  $p_R$  quedan relacionadas mediante:

$$\begin{aligned} x_L &= x_R + d \\ y_L &= y_R \end{aligned} \quad (10.1)$$

Empleando triángulos semejantes para los siguientes:  $\triangle PO_L O_R$ ,  $\triangle p_R O_R O_L$  y  $\triangle p_L O_L O_R$ , se tiene que:



$$d = \frac{f}{Z}B \quad (10.2)$$

Entonces, se tiene la relación entre  $d$  y  $Z$ :

$$d \propto \frac{1}{Z} \quad (10.3)$$

Basado en la ecuación 10.3 se puede recuperar, la profundidad de cada píxel en cada una de las imágenes.

### 10.2.2. Resultados

La figura 10.7 muestra el mapa de disparidad obtenido para un par de imágenes procesadas (Figura 10.6) con el conjunto correspondiente a la escultura de *Moises*.

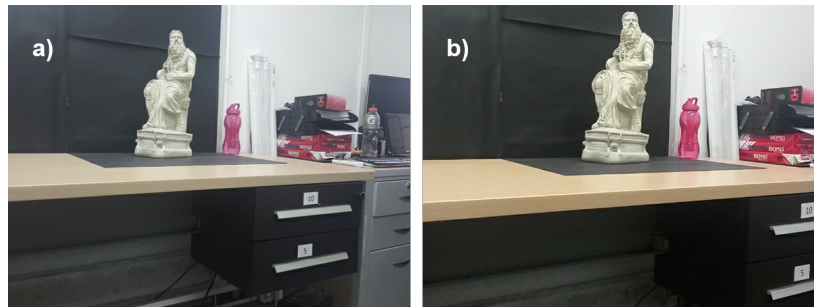


Figura 10.6: Imágenes procesadas. a) Imagen cámara izquierda b) Imagen cámara derecha

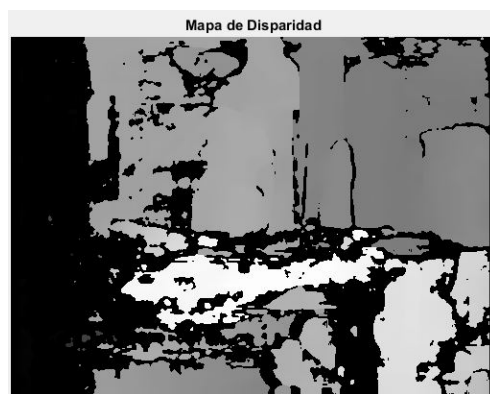


Figura 10.7: Mapa de Disparidad.

## 10.3. Reconstrucción de la Escena

La reconstrucción de la escena se obtiene a partir de la nube de puntos resultante de la triangulación, de esta manera es posible visualizar gráficamente el objeto de interés.

### 10.3.1. Procedimiento

La reconstrucción 3D se logra a partir del mapa de disparidad, y los datos de entrada son los parámetros obtenidos en el proceso de Calibración que se nombró anteriormente. Se definió la zona de interés para evitar ruido y datos innecesarios alrededor del objeto.

El proceso de reconstrucción tridimensional puede dividirse en dos fases:

1. Procesamiento digital de las imágenes
2. Visualización de la reconstrucción

La primera fase involucra la adquisición de las imágenes y posteriormente la búsqueda de correspondencia de puntos en las imágenes. La segunda es la reconstrucción a partir de las imágenes y la visualización del objeto.

### 10.3.2. Resultados

La nube de puntos obtenida con el teléfono inteligente a partir de pares de las imágenes de la escultura de *Moises* capturadas en diferentes posiciones contiene alrededor de 188738 puntos después de eliminar todos los puntos atípicos que generaban ruido en el modelo. La figura 10.8 muestra la escena obtenida a partir de dos imágenes capturadas por las cámaras del teléfono móvil seleccionado.

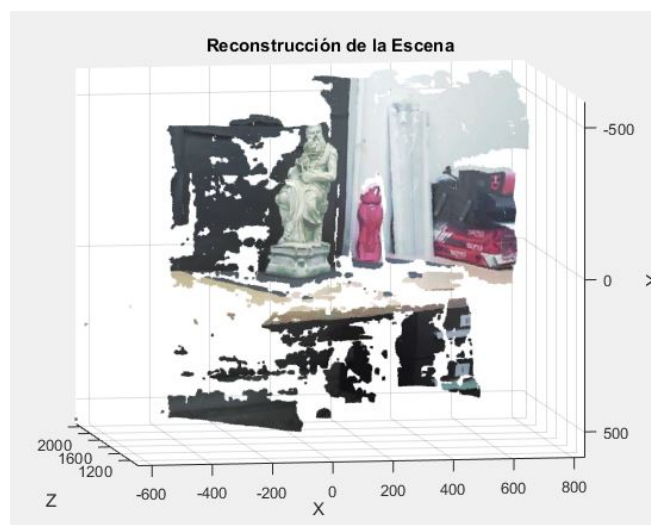


Figura 10.8: Reconstrucción de la escena

En la figura 10.8 se observa ruido, generado por el fondo y los objetos alrededor de la escultura, lo que produce errores en la reconstrucción. Por esta razón, se realiza una selección de la zona de interés donde se busca excluir la mayor parte de ruido. (figura 10.9)

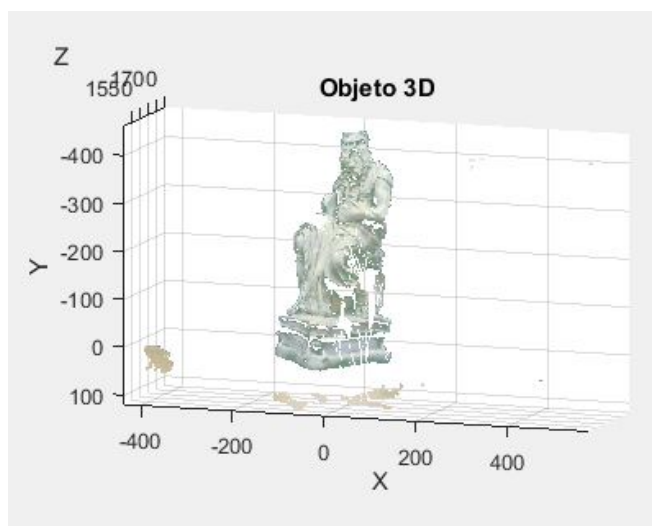


Figura 10.9: Reconstrucción escultura de Moises

## 10.4. Evaluación de la nube de puntos

Para mostrar el desempeño de la solución propuesta es necesario ponerlas a prueba bajo diversos escenarios, es por esto que se ha diseñado un experimento que consiste en dos fases que se explican a continuación:

Un análisis visual del modelo permitirá una evaluación global de la calidad del modelo, pero es importante realizar un análisis cuantitativo, para lo cual se utilizó el Software *Cloud Compare*, que permitió la manipulación de nubes de puntos 3D; además se realizó la comparación entre la nube de puntos y el modelo obtenido con el Kinect.

Aunque las imágenes obtenidas con el teléfono inteligente estén calibradas y rectificadas, las nubes de puntos obtenidas aparecen en posiciones, orientación y escala arbitrarias en el espacio tridimensional. Por esta razón, la nube de puntos es escalada manualmente usando el software *Cloud Compare* y con el mismo se alinea la nube de puntos del teléfono y el modelo del Kinect.

### ■ Prueba I

Uno de los aspectos que influye en la precisión de un modelo 3D es el valor del error de la nube de puntos que lo define, que es la diferencia entre el valor medido y el valor real (del mundo), siendo mayor cuanto más grande sea la

diferencia. A partir de lo anterior se utilizó un conjunto de imágenes de objetos geométricos sencillos con diferentes texturas de color, la nube de puntos obtenida está compuesta por 178284 después de excluir los datos atípicos de la escena.

Para esta fase se pretendió comparar las medidas obtenidas en la nube de puntos con la medida del mundo (real), con la intención de obtener distancias que permitan calcular un error cuantitativo. La Figura 10.10 muestra la imagen de los objetos y el cuadro 10.2 relaciona las distancias tomadas del objeto reconstruido y el objeto real.

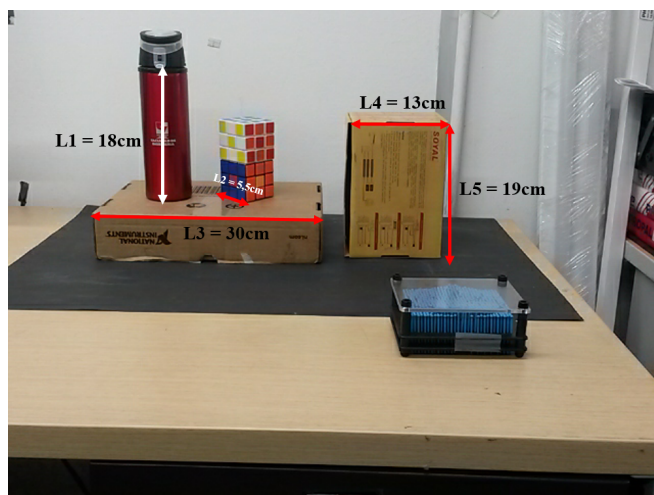


Figura 10.10: Dimensiones reales de los objetos

No.	Nube de Puntos (cm)	Objeto (cm)	Error (cm)	E. Relativo (cm)
L1	18.47	18	0.47	2.61 %
L2	5.70	5.5	0.2	3.6 %
L3	30.87	30	0.87	2.9 %
L4	14.21	13	1.21	9.3 %
L5	18.83	19	-0.17	0.89 %

Cuadro 10.2: Muestra de distancias

El cuadro 10.2 relaciona las distancias tomadas en la nube de puntos y el objeto real, indicando los errores relativos y absolutos entre cada una de las medidas. El error promedio calculado es  $\bar{X} = 0.51cm$ .

## ■ Prueba II

A partir de las imágenes de la escultura de *Moises*, se obtiene un nube de puntos compuesta por 188738 puntos. La alineación de la nube al modelo obtenido con el Kinect se realiza a partir de puntos característicos definidos manualmente,

la figura 10.11 muestra los modelos en el mismo sistema de referencia y puntos que fueron seleccionados para realizar la transformación.

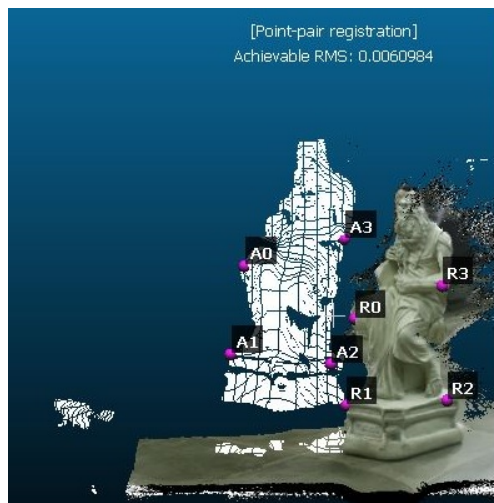


Figura 10.11: Alineación nube de puntos con modelo del Kinect

El cuadro 10.3 indica las coordenadas de los puntos seleccionados para la alineación. Este proceso tiende a presentar errores debido a que la selección de los puntos es manual; por esta razón, el cuadro 10.4 indica el error de alineación obtenido para cada punto, del cual se deduce;  $\tilde{X} = -0.12cm$  como error promedio y una desviación estándar de  $\sigma = 0.054cm$ .

Equipo	Punto	X	Y	Z
Kinect	AO	70.14	-140.16	1587.24
	A1	70.11	-140.32	1587.27
	A2	70.29	-140.35	1587.18
	A3	70.32	-140.11	1587.2
Nube de Puntos	R0	70.34	-140.26	1587.21
	R1	70.33	-140.42	1587.22
	R2	70.51	-140.41	1587.21
	R3	70.51	-140.19	1587.25

Cuadro 10.3: Puntos seleccionados para la alineación

Punto	E. Absoluto (cm)	E. Relativo
AO - R0	-0.07	-0.35 %
A1 - R1	-0.07	-0.38 %
A2- R2	-0.19	-0.35 %
A3 - R3	-0.16	-0.33 %

Cuadro 10.4: Error calculado en los puntos de referencia

La figura 10.12 muestra un histograma donde se representan la cantidad de puntos identificados y las distancias entre ellos. Cabe resaltar que no existe correspondencia entre el color y la distancia. El diagrama representa distancias negativas para los colores fríos y distancias positivas para los colores cálidos. De esta manera, se puede evidenciar que 7000 puntos de la nube y el modelo del kinect están a una distancia en el rango de -0.04cm a 0cm y alrededor de 4500 puntos están en el rango de 0cm a 0.04cm.

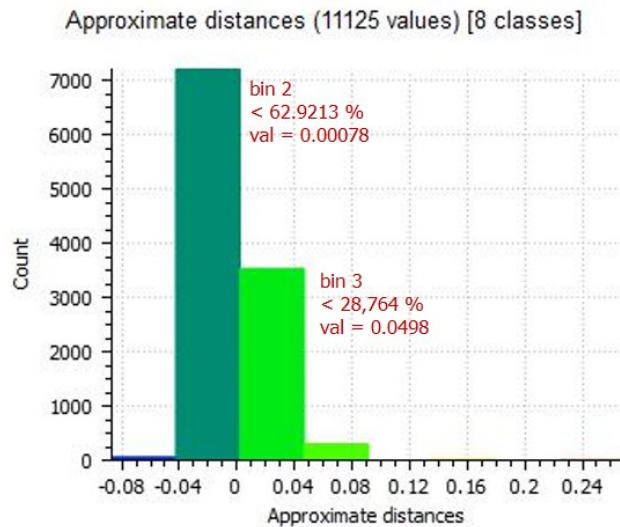


Figura 10.12: Representación gráfica de distancias entre nube de puntos y Kinect

Las imágenes capturadas de la escultura del museo arqueológico fueron post procesadas de la misma manera como se explica en el proceso anterior. La nube de puntos obtenida con las imágenes está compuesta por 206745 puntos, la cual fue alineada al modelo reconstruido con el Kinect de la misma escultura (ver figura 10.13).

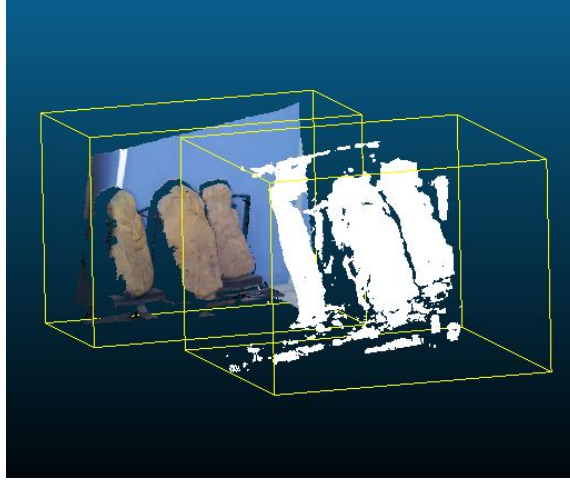


Figura 10.13: Alineación nube de puntos con modelo de la escultura

A diferencia del caso anterior, no se realizó el registro con la selección manual de puntos bien definidos. Esta vez el registro de la nube de puntos y el modelo de la escultura se realizaron a partir del algoritmo ICP, con el que se obtuvo la siguiente matriz de transformación 10.4:

$$\begin{bmatrix} 0.596 & -0.065 & -0.095 & 189.197 \\ 0.082 & 0.592 & 0.105 & -108.618 \\ 0.082 & -0.116 & 0.590 & 359.000 \\ 0.000 & 0.000 & 0.000 & 1.000 \end{bmatrix} \quad (10.4)$$

Una vez alineadas, se realiza el cálculo de distancias entre ambas superficies, la figura 10.14 indica el diagrama de barras resultante, que representa de colores fríos a colores cálidos las distancias negativas y positivas, respectivamente.

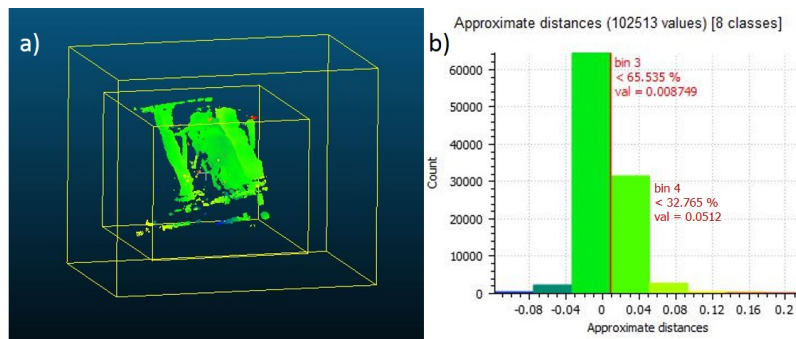


Figura 10.14: Distancia entre las dos superficies

De acuerdo con la gráfica, alrededor de 60000 mil puntos de las superficies

están a una distancia en el rango de  $-0.05\text{m}$  a  $0.01\text{m}$ , 30000 puntos entre  $0.01\text{m}$  a  $0.04\text{m}$ , 2000 puntos entre  $0.04\text{m}$  a  $0.09\text{m}$  y los puntos restantes están en el rango entre  $-0.11\text{m}$  a  $-0.05\text{m}$ .

Los dos primeros conjuntos de puntos nombrados anteriormente, representan la mayor cantidad puntos emparejados; con un  $65.53\%$  y  $32.76\%$  respectivamente, por tal razón; el porcentaje de alineación total del algoritmo es de  $98.29\%$ .

## 10.5. Limitaciones

- Un inconveniente del método desarrollado es que para ser utilizado es necesario realizar siempre la calibración de las cámaras antes de procesar las imágenes para reconstruir un objeto deseado. Siempre y cuando no se modifique la ubicación de los teléfonos inteligentes se podrán utilizar los mismos datos de la calibración inicial y capturar cuantas imágenes se desee de objetos de interés. En caso contrario, los valores de posición y orientación de la cámara no concidirán y dificultarán las fases posteriores.
- El patrón de calibración utilizado está construido por cuadrados de pequeña dimensión, debido a que el objeto de reconstrucción no es de gran tamaño. Hay que tener en cuenta que debe haber concordancia con la distancia de captura del patrón de calibración y la distancia de captura del objeto de interés.
- A pesar de contar con un ambiente controlado para la captura de las imágenes, los resultados obtenidos en la reconstrucción del objeto presentan ruido y datos que no proporcionan información valiosa para el modelo 3D. Es necesario, antes de fijar un resultado final, delimitar la nube de puntos sólo con el objeto de interés para evitar procesar datos atípicos.
- La generación del mapa de profundidad presentó dificultad en situaciones donde los valores de pixel son constantes y/o donde no hay cambio de texturas en el objeto de interés, el fondo u otro elemento que hiciera parte de la escena. Por esta razón es importante evitar objetos donde el algoritmo identifique incorrectamente un sector de una imagen en otra que no corresponde.
- La velocidad de la rutina de programación se relaciona directamente con la cantidad de imágenes que se deben procesar, permitiendo trabajar con una cantidad  $n$  de imágenes, difiriendo el tiempo de procesamiento de una iteración a otra.



# Capítulo 11

## Conclusiones

Este trabajo de grado proporciona una metodología para generar y evaluar un modelo tridimensional empleando imágenes capturadas con la cámara de teléfonos inteligentes, de modo que pueda ser usado por entidades de gestión, planeación y reconstrucción de objetos. La realización de éste tuvo interés especial en la aplicación de sensores de fácil acceso para la generación de información del terreno para evaluar la precisión del modelo tridimensional obtenido, desde lo cual se concluye lo siguiente:

- Se evaluó la precisión del modelo geométrico tridimensional obtenido mediante procesos fotogramétricos usando imágenes de una escultura capturadas con teléfonos inteligentes. Esta evaluación se llevó a cabo mediante el software *Cloud Compare* y una plataforma experimental. La evaluación está conformada por dos etapas de prueba: la primera con imágenes geométricas básicas y la segunda con imágenes de una escultura, comparadas con la medida real y con un modelo referencia obtenido con el Kinect alcanzando un error promedio de  $0.51cm$  y  $-0.12m$  respectivamente.
- Se realizó una revisión bibliográfica a través de la cual se estableció el fundamento conceptual, las teorías empleadas y el estado del arte de este trabajo de grado en torno a los temas de correspondencia entre imágenes e información de profundidad.
- El equipo hardware seleccionado como elemento de validación fue el Kinect, que permite visualizar elementos como: profundidad, tamaño y color hasta un metro de distancia, evitando la luz solar que impide el reconocimiento de los objetos.
- El diseño de la plataforma para calibración estéreo, permitió adaptar dos celulares inteligentes a una distancia determinada, propiciando la captura de las imágenes del patrón de calibración y de la escultura.

- La rutina de programación proporcionó entre sus procesos más significativos, a partir de un par de imágenes rectificadas de la escena capturadas con un desplazamiento horizontal en la plataforma de calibración, información de profundidad mediante el cálculo de disparidad.
- Se diseñó un protocolo de prueba conformado por imágenes de figuras geométricas básicas como: cubos y cilindros, permitieron visualizar la reconstrucción de la escena y verificar cuantitativamente las distancias definidas de los objetos.
- La precisión de los parámetros de calibración, permite obtener a partir de la información de profundidad, una correcta reconstrucción tridimensional, visualizando a partir de imágenes 2D un modelo 3D.
- Las imágenes fueron capturadas en un ambiente controlado, donde se mantienen las mismas condiciones de iluminación para evitar errores en la generación del mapa de disparidad y datos posteriores; sin embargo se presentan diferencias de iluminación entre las imágenes de una cámara a otra.
- La generación de una nube de puntos usando la cámara de los teléfonos inteligentes como sensor principal es una alternativa eficiente y efectiva en cuanto al tiempo de adquisición y procesamiento, resultando ser una propuesta novedosa para visualizar una escena tridimensional en un ambiente virtual.

## Capítulo 12

# Trabajos Futuros

Se considera que una futura línea de trabajo debería abordar el siguiente tema:

La nube de puntos obtenida con los telefonos inteligentes no siempre se visualiza con información de color, ya que cuando se pretende mostrar el color del objeto de interés es necesaria una imagen de la cámara para obtenerlo. Por esta razón cuando se pretende unir varias nubes de puntos de diferentes estereo imágenes, la reconstrucción no se puede visualizar con color, debido a que el tamaño y el volumen de la nube aumenta y la información de la imagen no alcanza a registrar el color. Aunque esto no es una limitación para analizar los resultados, el ejercicio de lograr una reconstrucción 3D a partir de varias uniones de nubes de puntos recuperando el color resulta un tema interesante de abordar.

# Bibliografía

- Agarwal, S., N. Snavely, I. Simon, S. M. Seitz, and R. Szeliski (2011), Building Rome in a Day, *Communications of the ACM*, 54, 105–112, doi:10.1145/2001269.
- Almagro, A. (2009), De la fotogrametría a la infografía. Un proceso informatizado de documentación.
- Bay, H., T. Tuytelaars, and L. V. Gool (2006), SURF: Speeded up robust features, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3951 LNCS, 404–417, doi:10.1007/11744023\_32.
- Bay, H., A. Ess, T. Tuytelaars, and L. V. Gool (2008), Speeded-Up Robust Features ( SURF ), *Elsevier*, (September), 2–14.
- Bejarano, J. D. M. (2013), Técnicas de fotogrametría y visión por computador para el modelado 3D de estructuras geomorfológicas dinámicas, Ph.D. thesis.
- Chen, L., S. Betschart, and A. Blaylock (2015), Projeto Redentor, *Tech. rep.*
- Chiabrando, F., F. Nex, D. Piatti, and F. Rinaudo (2011), UAV and RPV systems for photogrammetric surveys in archaeological areas: two tests in the Piedmont region (Italy), *Journal of Archaeological Science*, 38(3), 697–710, doi:10.1016/j.jas.2010.10.022.
- Espín, M. A., and F. A. Guano (2013), Diseño de un ambiente virtual mediante el empleo de procesos fotogramétricos terrestres de corto alcance, del conjunto arquitectónico de la ciudad de Conocoto, Ph.D. thesis.
- Esteban, I., J. Dijk, and F. Groen (2010), FIT3D toolbox: multiple view geometry and 3D reconstruction for Matlab, 7835, 78,350J–78,350J–9, doi:10.1117/12.864112.
- Fernandez, W. D., J. D. Pacateque, M. S. Puerto, M. I. Balaguera, and F. Reyes (2015), Reconstrucción digital de mezclas asfálticas basadas en imágenes de tomografía computarizada, *Ciencia e Ingeniería Neogranadina*, 25, 17–26.
- Furukawa, Y., and C. Hernández (2013), Multi-View Stereo: A Tutorial, *Foundations and Trends in Computer Graphics and Vision*, 9(1-2), 1–148.

- Ginovart, J., J. Toldrá, A. Costa, and S. Coll (2014), Close Range Photogrammetry and Constructive Characterization of Masonry Gothic Vaults, *Journal Of Construction*, 13(1), 47–55.
- Gonzalez, J. I. (2003), Estudio experimental de métodos de calibración y autocalibración de cámaras, Ph.D. thesis.
- Hartley, R., and A. Zisserman (2004), Introduction. A Tour of Multiple View Geometry, in *Multiple View Geometry in Computer Vision*, pp. 1–21.
- Hernández López, D. (2006), *Introducción a la fotogrametría digital*, 1–188 pp.
- Hu, S., J. Qiao, A. Zhang, and Q. Huang (2008), 3D Reconstruction From Image Sequence, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII, 559–562.
- Jauregui, L. (2006), Introducción a la fotogrametría, in *Introducción a la fotogrametría*, pp. 1–12.
- Jauregui, L., M. Jauregui, L. Chacón, and J. Vilchez (2013a), Determinación de los parámetros de orientación interior de una cámara digital no métrica., *Revista Ciencia e Ingenieria*, 34(1), 31–36.
- Jauregui, M., L. Jauregui, L. Chacón, and J. Vilchez (2013b), MODELADO Y VISUALIZACION TRIDIMENSIONAL DE LA ESTATUA DE MARÍA LIONZA, *Tech. rep.*
- Kim, J., S. Lee, H. Ahn, D. Seo, D. Seo, J. Lee, and C. Choi (2013), Accuracy evaluation of a smartphone-based technology for coastal monitoring, *Measurement*, 46(1), 233–248, doi:10.1016/j.measurement.2012.06.010.
- Kim, K. J. (2014), Can smartphones be specialists? Effects of specialization in mobile advertising, *Telematics and Informatics*, 31(4), 640–647, doi: 10.1016/j.tele.2013.12.003.
- Lerma, J. L. (2011), Líneas de investigación del grupo de investigación: fotogrametría, modelización y clasificación, *Tech. rep.*
- Lerma, J. L., M. Cabrelles, A. E. Seguí, and S. Navarro (2011), Aplicación de la fotogrametría terrestre al levantamiento de alzados de edificios singulares, *Instituto Andaluz del Patrimonio Historico*, 77, 127–129.
- Lin, G. Y., X. Chen, and W. G. Zhang (2010), A robust epipolar rectification method of stereo pairs, *2010 International Conference on Measuring Technology and Mechatronics Automation, ICMTMA 2010*, 1, 322–326, doi: 10.1109/ICMTMA.2010.220.
- Ma, Y., J. Kosecka, S. Soatto, S. Sastry, and J. Ko (2006), *An Invitation to 3-D Vision: From Images to Models*, 526 pp., doi:10.1007/978-0-387-21779-6.
- Mccann, S. (2014), 3D Reconstruction from Multiple Images, *Tech. rep.*

- Nájera, G. (2009), Realidad Aumentada en Interfaces Hombre Maquina, Ph.D. thesis, Instituto Politécnico Nacional.
- Olsson, C., F. Kahl, and M. Oskarsson (2006), The Registration Problem Revisited : Optimal Solutions From Points , Lines and Planes, *IEEE*, 6, 1–8.
- Pacheco, A., H. Bolívar, J. P. Espada, and R. G. Crespo (2014), Reconstruction of High Resolution 3D Objects from Incomplete Images and 3D Information, *International Journal of Artificial Intelligence and Interactive Multimedia*, 2, 7–16, doi:10.9781/ijimai.2014.261.
- Penna S, J., E. A. Cárdenas Q, and L. Y. Morales (2014), Aplicaciones de la estereoscopia en la reconstrucción tridimensional de zonas de terrenos, *XVI Simposio Internacional SELPER 2014*, pp. 1–16.
- Quintero, D., R. Villamizar, and E. Mendoza (2011), Reconstrucción tridimensional de deformaciones mecánicas superficiales en platinas , basada en la proyección de franjas., *Facultad Ingenieria, Universidad de Antioquia*, 59, 169–180.
- Revelo-Luna, D. A., F. D. Usama, and J. F. Flórez-Marulanda (2012), Reconstrucción 3D de escenas mediante un sistema de visión estéreo basado en extracción de características y desarrollado en OpenCV, *Ingenieria Universidad Nacional de Colombia*, pp. 485–500.
- Roncella, R., C. Re, and G. Forlani (2011), Performance Evaluation of a Structure and Motion Strategy in Architecture and Cultural Heritage, *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38, 285–292, doi:10.5194/isprsarchives-XXXVIII-5-W16-285-2011.
- Rosado, J. (2012), Reconstrucción de escenas 3D a partir de un conjunto de vistas tomadas desde una cámara, Ph.D. thesis.
- Sánchez, N., B. Arias, D. González, and J. Gómez (2004), Análisis aplicado de métodos de calibración de cámaras para usos fotogramétricos, *Congreso Nacional de Topografía y Cartografía*, pp. 1–15.
- Santamaría, J., and F. Sanz (2013), Modelado 3D de terrenos a partir de fotografías, *VII Congreso Ibérico de Agroingeniería y Ciencias Hortícolas*, C0539, 6.
- Sun, J., X. Chen, Z. Gong, Z. Liu, and Y. Zhao (2015), Accurate camera calibration with distortion models using sphere images, *Optics & Laser Technology*, 65, 83–87, doi:10.1016/j.optlastec.2014.07.009.
- Szeliski, R. (2010), *Computer Vision : Algorithms and Applications*, vol. 5, 832 pp., doi:10.1007/978-1-84882-935-0.

- To, T., D. Nguyen, and G. Tran (2015), AUTOMATED 3D ARCHITECTURE RECONSTRUCTION FROM PHOTOGRAMMETRIC STRUCTURE AND MOTION : A CASE STUDY OF THE "ONE PILLAR" PAGODA , HANOI , VIETNAM, *36th International Symposium on Remote Sensing of Environment, XL*(May), 11–15, doi:10.5194/isprsarchives-XL-7-W3-1425-2015.
- Vite Silva, I. (2008), Reconstrucción Tridimensional de Objetos Mediante Técnicas Evolutivas, Ph.D. thesis, Universidad Zacatenco.
- Vozmediano, D. (2006), Fotogrametría digital aplicada a la obtención de ortofotos y modelos digitales de entidades patrimoniales, Ph.D. thesis.
- Westoby, M., J. Brasington, N. Glasser, M. Hambrey, and J. Reynolds (2012), Structure-from-Motion photogrammetry: A low-cost, effective tool for geoscience applications, *Geomorphology*, *179*, 300–314, doi:10.1016/j.geomorph.2012.08.021.
- Yuan-yuan, T., L. Si-yang, and T. Qing-chang (2012), Application of Detecting Part's Size Online Based on Machine Vision, *Energy Procedia*, *16*, 1948–1956, doi:10.1016/j.egypro.2012.01.298.
- Yun, M., J. Kim, D. Seo, J. Lee, and C. Choi (2012), APPLICATION POSSIBILITY OF SMARTPHONE AS PAYLOAD FOR PHOTOGRAMMETRIC UAV SYSTEM, *XXII ISPRS Congress, XXXIX-B4*, 349–352.
- Zheng, J., L. Zhang, and X. Du (2009), Accurate 3D Target Positioning in Close Range Photogrammetry with Implicit Image Correction, *Chinese Journal of Aeronautics*, *22*(6), 649–657, doi:10.1016/S1000-9361(08)60154-5.